HUMAN FACTORS, 1990, 32(6), 637–649

Comprehension and Memory of Instruction Manual Warnings: Conspicuous Print and Pictorial Icons

STEPHEN L. YOUNG, *Rice University, Houston, Texas, and* MICHAEL S. WOGALTER,[1]
Rensselaer Polytechnic Institute, Troy, New York

Two experiments examined the effects of increasing the noticeability of instruction manual warnings on subsequent comprehension and memory performance. Participants read one of four instruction manuals for a gas-powered electric generator (Experiment 1) or a natural-gas oven (Experiment 2) on the assumption that they would later operate the equipment. The appearance of eight different warning messages in the manuals was altered in two ways: (1) the verbal messages were printed either in conspicuous print (larger text with color highlighting) or in plain print (same as the other text), and (2) either the verbal warning messages were accompanied by compatible pictorial icons or the icons were absent. Results showed that participants who received the conspicuous print, icons present manual better comprehended and recalled the verbal warning messages (Experiments 1 and 2) and better identified the semantic meaning of the icons (Experiment 1) than did participants who received the other three manuals. Implications for the design of instruction manual warnings are discussed.

INTRODUCTION

Most new electrical and gas-powered equipment is accompanied by an instruction manual describing the product's correct setup and operation. Usually the manuals also contain important safety information in the form of warnings. The purpose of warnings is to communicate the nature of potential hazards associated with the product and the procedures for using the equipment so that injury to the person and damage to the machine can be avoided.

[1] Requests for reprints should be sent to Michael S. Wogalter, Psychology Department, Rensselaer Polytechnic Institute, Troy, NY 12180.

Because consumers may not always have the manual available when they use a hazardous product, it seems vital that the warnings be designed not only to facilitate comprehension the first time the user examines the material but also to enhance recall of the hazard information when the product is used at a later time. Surprisingly, however, reviews of the literature indicate that research on the factors that influence warning comprehension and memory is scarce (DeJoy, 1989; Lehto and Miller, 1986). The few existing studies have shown no effect or only small effects of various format manipulations (Desaulniers, 1987; Otsubo, 1988; Rothstein,

1985; Strawbridge, 1986; Zlotnik, 1982). The purpose of the current research is to examine two factors that might influence comprehension and memory of warnings: (1) increased conspicuity of printed text and (2) the presence of pictorial icons.

One way to increase warning comprehension and memory is to improve the likelihood that they are noticed and read initially. Increasing warning noticeability usually involves making the warning text more conspicuous (salient) than other textual material in a manual. Indeed, virtually every set of published guidelines on warnings emphasizes the characteristic of conspicuousness or salience by recommending that warnings should stand out from a noisy background (Cunitz, 1981; Peters, 1984; Wogalter et al., 1987). One means of increasing text conspicuity is to highlight the print (e.g., by adding color or by increasing its size and boldness). Several recent studies have noted positive effects of highlighting on reading warnings (Strawbridge, 1986) and complying with them (Zlotnik, 1982). However, no published research demonstrates greater comprehension and memory for highlighted warning text than for unhighlighted text. Thus one of the factors examined in the present research concerned the effect of warning text conspicuity on comprehension and memory.

Another possible way to increase comprehension and memory of warnings is to include icons (symbols or pictograms). Most warnings guidelines (FMC, 1985; Westinghouse, 1981; Wogalter et al., 1987) recommend that warnings include a descriptive pictorial to facilitate communication of hazards. Dorris and Purswell (1978) suggested that icons might be more rapidly recognized and more effective at conveying hazard information than would a verbal message. Robinett and Hughes (1984) pointed out that because of the complex nature of many hazards, icons cannot realistically exist by themselves as a means of communicating hazard information. They noted that in most situations, icons should be paired with verbal warning messages.

Research on the effectiveness of pictorials in warnings has been surprisingly sparse, though two recent studies have cast doubt on the usefulness of pictorials in warnings. Friedmann (1988) found no effect of pictorials on warning compliance, and Otsubo (1988) found no significant differences among warnings with words only, pictographs only, and words plus pictographs on noticing, reading, complying with, and recalling the warning message. Most other warning icon research has dealt with people's ability to identify icons (e.g., Collins, Lerner, and Pierman, 1982; Laux, Mayer, and Thompson, 1989; Mayer and Laux, 1989). Because pictorials are often mentioned as an important component of effective warning design, and because research has failed to demonstrate their effectiveness (beyond their ability to be identified), the present research reexamined the usefulness of icons on comprehension and memory of the associated warning message.

The presence of icons could facilitate memory for warnings in another way. By pairing icons and verbal warnings, the two may become associated in memory, and then, at reexposure, the icon might facilitate retrieval of the hazard information by cuing the warning message. To investigate this, an icon identification test examined whether participants could recall the associated hazard information given the icons as cues. In addition, another icon memory test—icon recognition—was included in Experiment 1 to determine whether participants could distinguish the icons that were presented in the manuals from other (distractor) icons.

In warnings research ethical considerations prohibit testing of warning effectiveness in situations where actual danger is involved. The present research used a meth-

odology that avoided this problem by constructing a situation that was a ruse. The procedure involved two phases. Participants first performed a set of computer tasks with an instruction manual *present*. After these tasks were completed, the experimenter described the second task as a test of people's ability to operate complex equipment with instructions *absent*. They were given a manual, asked to examine it, and told that they would have to operate the equipment later (though no one actually did). Thus participants were led to believe that they would have to operate a dangerous piece of equipment but were instead given comprehension and memory tests.

At this point a distinction should be made between warning comprehension and memory as they are used in the present context. *Comprehension* refers to understanding the safety information in the warnings, and *memory* refers to recall and recognition of specific information (e.g., particular wording, icons) in the warnings. All tests were given while the manuals were no longer being viewed, and thus all could be technically described as tests of memory. However, the tests and the methods of scoring the tests could potentially reveal different kinds of information. For example, one test assessed knowledge of the verbal message content of the warnings. Performance on this test was measured by grading the tests in two ways, using liberal and strict criteria. Liberal scoring assessed participants' general understanding or *comprehension* of the safety information in the warnings. Strict scoring assessed *memory* of specific details in the particular warnings.

Two experiments are reported which examined whether increasing the salience of warning messages in a set of proceduralized instructions would improve comprehension and memory. Participants examined one of four instruction manuals for a gas-powered electric generator (Experiment 1) or a natural-gas oven (Experiment 2) which differed with respect to two factors: (1) the verbal messages were printed either in conspicuous print (larger print and in a different font style with color highlighting) or in plain print (same as the other text), and (2) either the verbal warning messages were accompanied by related pictorial icons or the icons were absent. It was hypothesized that warnings printed conspicuously and paired with pictorial icons would produce greater comprehension and memory than would warnings in plain print and/or with pictorial icons absent.

EXPERIMENT 1

Experiment 1 used instruction manuals for a gas-powered electric generator. This particular piece of equipment was chosen for three reasons: (1) for most persons, it is an unfamiliar piece of equipment; (2) it poses some risk of injury; and (3) given that the generator is designed for use in a power outage, it would have to be operated safely in conditions in which it is impractical or impossible to reference an instruction manual (e.g., in rain or absence of light—situations where memory of the warnings may be necessary).

Comprehension and memory of the warnings were assessed using three tests. The first test measured participants' recall of the warning message content. The second and third tests measured comprehension and memory of the icons. One icon test assessed participants' ability to recognize the target icons embedded in a larger set of distractor icons. The other icon test measured participants' ability to verbally identify the intended meaning of the icons. Different kinds of knowledge are assessed by these three tests, and they potentially could reveal different effects on warning comprehension and memory.

Method

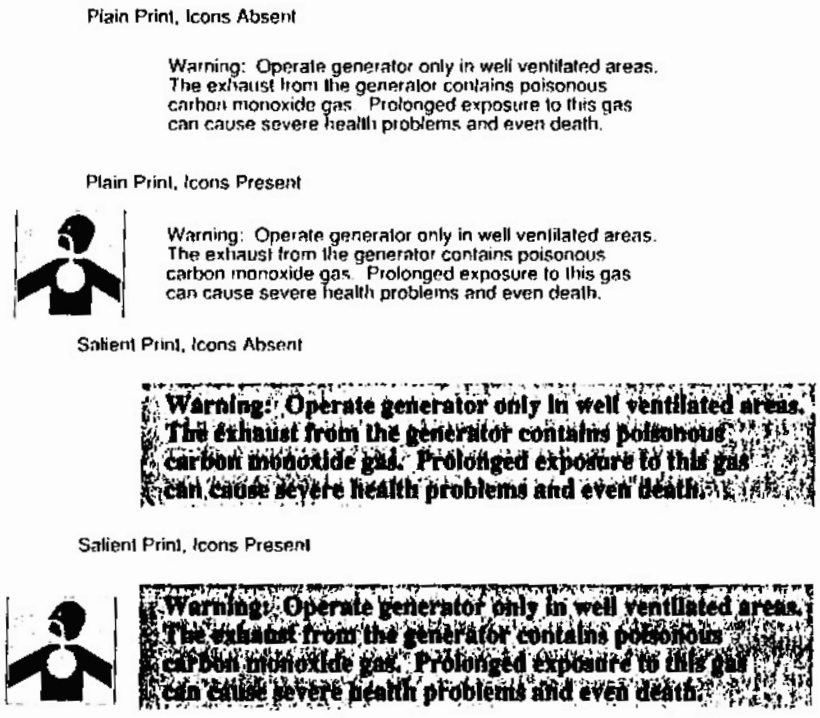
Subjects. Sixty-four undergraduate students from the University of Richmond par-

ticipated in the experiment for credit in an introductory psychology course.

Materials. The basic 10-page instruction manual described the operation and maintenance of a gas-powered electric generator. The manual was adapted from several existing manufacturers' manuals and had the appearance of an actual owner's manual. The text of the manuals was printed in 12-point Helvetica font (except for the headings, which were in 12-point Helvetica bold font). It contained sections on the generator's operation and maintenance as well as detailed drawings of the generator.

Each manual contained eight verbal warning messages. The warnings were printed either in conspicuous print or in plain print. Conspicuous print was 18-point Times font covered with transparent orange fluorescent highlighting. Plain print warnings had the same appearance as the other textual print in the manual. The warnings were accompanied by compatible icons, or the icons were absent. The icons were obtained from a technical report by Collins et al. (1982) in which a large number of pictorial icons were evaluated on comprehension. Nine icons were used: one each for seven warnings and two for a gas explosion warning. Four instruction manuals were produced: (1) conspicuous print, icons present, (2) conspicuous print, icons absent, (3) plain print, icons present, and (4) plain print, icons absent. The warning manipulations are illustrated in Figure 1.

Three tests were used to assess participants' comprehension and memory of the



Note: Shading represents orange highlighting.

Figure 1. *Four variations of an example warning.*

warning messages and icons: message content, icon recognition, and icon identification. The message content test consisted of 10 short-answer questions. One question was developed for each of the eight warning messages. Each dealt with some aspect of the semantic content of the warning message (e.g., "In what kind of environment must generator maintenance be done and why?"). Two of the test questions concerned operational aspects of the generator and were not analyzed. Space was provided under each question for responses.

The icon recognition test consisted of a set of 36 randomly ordered icons. Nine were target icons that also appeared in manuals with icons present. The test set also included nine icons (*target-similar distractors*) that were similar in meaning to the target icons but which were graphically different (Collins et al., 1982). The other 18 icons (*target-dissimilar distractors*) consisted of nine pairs of icons that were not used in any instruction manual. The icons in each pair of target-dissimilar distractors were similar in meaning but were graphically different. A numbered sheet was provided for responses.

In the icon identification test, the nine icons used in the manuals were randomly ordered on a blank sheet. Space for responses was provided next to each icon.

Procedure. Participants were given one of the four generator instruction manuals and were told that they would have to know how to operate the generator from memory later in the session. The notion that participants would have to operate the generator was enhanced by having them engage in a series of computer interaction tasks before taking part in the generator manual task. Participants first performed the computer tasks while viewing a booklet of instructions in one room; then they were led to another room to study the generator manuals. The impression given to the participants was that the first set of tasks tested their ability to operate a piece of complicated equipment with the instructions present, and that the next (generator manual) task would test their ability to operate a piece of equipment with the instructions absent (i.e., from memory).

After having 4 min to examine the manual, all participants were told that they were in the "control" group and would not be operating the generator. Instead they were given the message content, icon recognition, and icon identification tests.

In the message content test, participants answered questions on the hazard information conveyed by the warnings in the manual. Participants were instructed to answer the questions as specifically as they could and to guess if necessary. In the icon recognition test, the participants attempted to recognize the target icons in a set of distractor icons. For each of the 36 icons in the test, participants wrote a *Y* (meaning yes, the icon was in the manual) or an *N* (meaning no, the icon was not in the manual) in the corresponding numbered blank on a response sheet. Participants also indicated their confidence in their answers by writing a 1 (guessed the answer), 2 (fairly sure of the answer), or 3 (very sure of the answer) next to their Y or N response. In the icon identification test, participants were instructed to describe the meanings of the icons. They were told that their descriptions should relate specifically to the warnings in the generator instruction manual and were told to guess if necessary.

Two aspects of the procedure warrant comment. The first concerns the fact that the tests were given in a fixed rather than a counterbalanced order. The message content test was always given first to avoid possible carryover effects from the other two tests. The other two tests were given in an order that had the least potential for carryover. The second comment concerns the fact that all participants were required to take both icon-related tests re-

gardless of whether they had been exposed to a manual containing icons. The reason for this arrangement was to keep all participants occupied on a task and to hold test onset times constant. In order to measure baseline icon understandability, participants in all conditions were given identical instructions presented in such a way that participants who had not been exposed to icons could give reasonable responses.

Prior to the initial procedure, participants were given a consent form to sign which stated that they were free to discontinue their participation from the experiment without penalty. On completion of the experimental tasks, the experimenter debriefed participants on the procedures and purposes of the study.

Results

All response sheets were coded, randomly ordered, and then scored by the first author without knowledge of conditions (i.e., blind). The data were analyzed using a set of planned comparisons that contrasted performance of the conspicuous print, icons present warnings with performance of the other three conditions. In addition, 2 (conspicuous vs. plain print) × 2 (presence vs. absence of icons) between-subjects analyses of variance (ANOVAs) were used to examine whether the print and icon variables produce linear and/or interactive effects.

Message content. Responses to the message content questions were scored twice according to liberal (comprehension) and strict (memory) criteria. To be scored correct under the liberal criterion, responses had to convey the general meaning of the warning message (e.g., something about gas and breathing problems for the example in Figure 1). To be scored correctly under the strict criterion, responses had to include certain key words and clearly demonstrate recall of specific information from the manuals' warnings (e.g., mention of the words "ventilation," "carbon monoxide," and/or "exhaust" for the example in Figure 1). Correct answers for both criteria were scored as 1 and incorrect answers were scored as 0.

TABLE 1

Mean Proportion Correct as a Function of Warning Format

	Conspicuous Print		*Plain Print*	
	Icons Present	*Icons Absent*	*Icons Present*	*Icons Absent*
Experiment 1 (Gas-Powered Electric Generator)				
Message Content				
Liberal	0.75	0.56	0.58	0.53
Strict	0.47	0.27	0.32	0.26
Icon Identification				
Liberal	0.93	0.81	0.86	0.85
Strict	0.48	0.21	0.26	0.10
Experiment 2 (Natural Gas Oven)				
Message Content				
Liberal	0.81	0.66	0.64	0.56
Strict	0.67	0.48	0.42	0.39
Icon Identification				
Liberal	0.81	0.84	0.83	0.73
Strict	0.26	0.18	0.16	0.14

Table 1 shows the mean proportion correct for the message content answers using the liberal and strict criteria. The conspicuous print, icons present condition produced greater performance than the other three conditions. The liberal scores are higher than the strict scores, but the pattern of means is similar for both measures.

Using the liberal scores, planned comparisons showed that comprehension of conspicuous print, icons present warnings was significantly greater than that of the other three warning conditions: conspicuous print, icons absent, $t(30) = 3.03$, $p < 0.01$; plain print, icons present, $t(30) = 2.26$, $p < 0.04$; and plain print, icons absent, $t(30) = 4.12$, $p < 0.001$.

Using the strict scores, conspicuous print, icons present warnings produced significantly greater memorial recall than did the conspicuous print, icons absent, $t(30) = 3.34$, $p = 0.01$, and plain print, icons absent warnings, $t(30) = 3.38$, $p < 0.01$. The difference between conspicuous print, icons present and plain print, icons present warnings failed to reach the criterion for significance, $t(30) = 1.95$, $p = 0.06$.

Overall analyses using a 2 (print) × 2 (icons) ANOVA on the liberal content scores showed significant main effects of print, $F(1,60) = 4.07$, $MS_e = 0.038$, $p < 0.05$, and icons, $F(1,60) = 6.28$, $p < 0.02$. Comprehension was greater with conspicuous print ($M = 0.65$) than with plain print ($M = 0.55$) and when icons were present ($M = 0.66$) than when they were absent ($M = 0.54$). The interaction was not significant, $F(1,60) = 1.92$, $p > 0.05$. The ANOVA on the strict content scores yielded a significant main effect of the icons, $F(1,60) = 7.53$, $MS_e = 0.038$, $p < 0.01$. Recall was higher when icons were present ($M = 0.40$) than when they were absent ($M = 0.26$). Neither the print main effect nor the interaction was significant, $F(1,60) = 2.67$, $p > 0.05$, and $F(1,60) = 2.07$, $p > 0.05$, respectively.

Icon recognition. Responses in the icon recognition test were analyzed in a number of ways. Responses to the target icons were scored as hits (1's for Y responses) or misses (0's for N responses). This test also contained a set of target-similar distractors (similar in meaning to the target icons) and a paired set of target-dissimilar distractors. Y responses to the distractor icons were scored as false alarms (1's for yes responses), except for one analysis in which the target-similar distractors were scored liberally as hits to capture recognition of the icons' meaning. In other analyses the confidence ratings were combined with the yes/no responses to produce a widened recognition confidence scale (*N3* assigned a 1, *N2* assigned a 2, *N1* assigned a 3, *Y1* assigned a 4, *Y2* assigned a 5, and *Y3* assigned a 6) ranging from 1 (very sure the icon was not seen before) to 6 (very sure the icon was seen before). This transformation is commonly used in recognition memory research because it is often more sensitive at finding differences between conditions than yes/no binary responses. In addition, two discrimination (sensitivity) measures were computed from the hit and false alarm scores.

The results showed the expected differences between the presence and absence of icons in the manuals for all of the aforementioned measures (p's < 0.05). However, no other effects (including relevant comparisons between the conspicuous print, icon present and plain print, icon present conditions) were significant using any of the measures (p's > 0.05).

Icon identification. The icon identification test was graded using both liberal and strict criteria. To be scored correct with the liberal criterion, responses had to convey the general meaning of the icon (comprehension). To be scored correct with the strict criterion, responses had to not only identify the meaning of the icon but also include specific information from the associated verbal warning

message (memory). Correct and incorrect responses were scored as 1's and 0's, respectively.

Table 1 shows that the conspicuous print, icons present warnings produced significantly better icon identification performance than did the other three warning conditions. Planned comparisons using the liberal scores showed that conspicuous print, icons present warnings produced significantly higher icon comprehension scores than did conspicuous print, icons absent, $t(30) = 3.25, p < 0.01$, and plain print, icons absent warnings, $t(30) = 2.30, p < 0.03$. The difference between conspicuous print, icons present and plain print, icons present warnings failed to reach the criterion for significance, $t(30) = 1.87, p < 0.08$. Comparisons using the strict scores showed that conspicuous print, icons present warnings produced significantly greater specific recall of the warning message than did the other three conditions: with conspicuous print, icons absent, $t(30) = 3.09, p < 0.01$, with plain print, icons present, $t(30) = 2.47, p < 0.02$, and with plain print, icons absent, $t(30) = 5.52, p < 0.001$.

Overall analyses using a 2 (print) × 2 (icons) ANOVA on the liberal scores produced a significant main effect of icons, $F(1,60) = 7.01, MS_e = 0.011, p < 0.02$. Performance was higher when icons were present ($M = 0.40$) than when they were absent ($M = 0.26$). Conspicuous print did not produce a main effect, $F(1,60) < 1.0$, but it did enter into an interaction with icons, $F(1,60) = 4.49, p < 0.05$. A Newman-Keuls post hoc test showed that conspicuous print warnings benefited from the addition of icons, but no such benefit was shown for the plain print warnings.

The ANOVA on the strict scores yielded significant main effects of print, $F(1,60) = 9.16$, $MS_e = 0.049, p < 0.004$, and icons, $F(1,60) = 15.10, p < 0.001$. Recall was greater with conspicuous print ($M = 0.34$) than with plain print ($M = 0.18$) and when icons were present ($M = 0.37$) than when they were absent ($M = 0.15$). The interaction was not significant, $F(1,60) = 1.01, p > 0.05$.

Reliability. Twenty response sheets from the message content and icon identification tests were rescored by a second judge to assess reliability of the scoring procedure. Selection of response sheets from each test was random with the constraint that each condition was represented equally (five from each condition). Correlations were then obtained between the original scores and the corresponding scores given by the second judge. For the message content test, the reliabilities were 0.87 and 0.88 for the strict and liberal scores, respectively. For the icon identification test, the strict and liberal scores showed respective reliabilities of 0.89 and 0.85.

Discussion

The results reveal that conspicuous print, icons present warnings were comprehended and remembered better than were the other three warning conditions. The presence of icons facilitated not only performance on the icons themselves but also understanding of the icon's associated warning message. The beneficial effect of conspicuous print was smaller than that of the presence of icons but, nevertheless, generally helped to produce greater message content and icon identification scores.

The icon recognition test failed to produce interesting results. As expected, icon recognition differed with respect to whether participants did or did not have icons in their manuals, but no other effects were noted.

EXPERIMENT 2

Experiment 2 examined whether the effects found in Experiment 1 would be found using a manual for a different product—namely, a natural-gas oven/range. The oven was chosen because this appliance posed the same kinds of hazards as did the generator (i.e., electric

shock, gas fire, and explosion), which meant that the same icons that were used for the generator could be used in the manual for the oven.

Method

Subjects. Sixty-four undergraduates from the University of Richmond (n = 52) and Rice University (n = 12) participated for credit in introductory psychology courses. Participants from both universities were equally distributed in the four conditions. None had participated in Experiment 1.

Materials. Four instruction manuals describing the operation and maintenance of a natural-gas oven were prepared. Like the generator manual, the oven manual was adapted from several existing manufacturers' manuals and had the appearance of an actual owner's manual, including detailed drawings. As in Experiment 1, the manuals differed with respect to the print and icon factors: (1) conspicuous print, icons present, (2) conspicuous print, icons absent, (3) plain print, icons present, and (4) plain print, icons absent. The icons were identical to those used in Experiment 1, except one of the two icons used for the gas explosion warning in Experiment 1 was dropped in Experiment 2. The verbal warning messages for the oven differed from those of the generator in specific details but were generally similar because both products had hazards in common. For example, the verbal message associated with the icon shown in Figure 1 was as follows:

> Warning: Power outages may cause the pilot lights to be extinguished, leaking gas into the air. This gas is poisonous and will cause harm if inhaled over a period of time. If you smell gas, turn off the gas and call your dealer immediately.

Two tests were used: a message content test and an icon identification test. The message content test consisted of eight short-answer questions, one for each of the eight warnings, which assessed some aspect of the semantic content of the warning message (e.g., "If you smell gas, what does it mean and what should you do about it?"). For the icon identification test, the eight icons were placed in random order on a response sheet. The icon recognition test was dropped because it failed to provide informative results in Experiment 1.

Procedure. Participants were given one of the four gas oven instruction manuals and told that they would be required to operate the oven from memory later in the session. Participants were given the same set of computer tasks as in Experiment 1 before receiving the oven manuals in order to enhance the impression that they would be actually operating an oven. Other aspects of the procedure were identical to those in Experiment 1, except that participants were given the two tests (i.e., message content and icon identification) in a counterbalanced order.

Results

The tests were scored as in Experiment 1. The proportion correct data for the liberally and strictly scored message content and icon identification tests are shown in the bottom half of Table 1.

Message content. As can be seen in Table 1, conspicuous print, icons present warnings produced greater comprehension and recall than did the other three warning conditions. Planned comparisons using the liberal (comprehension) scores confirmed that these differences were reliable: with conspicuous print, icons absent, $t(30) = 2.11$, $p < 0.05$; with plain print, icons present, $t(30) = 2.88$, $p < 0.01$; and with plain print, icons absent, $t(30) = 3.26$, $p < 0.01$. These effects were also significant using the strict (memory) scores: with conspicuous print, icons absent, $t(30) = 2.13$, $p < 0.05$; with plain print, icons present, $t(30) = 3.91$, $p < 0.001$; and with plain print, icons absent, $t(30) = 3.97$, $p < 0.001$.

Overall analyses using a 2 (print) × 2

(icons) ANOVA on the liberal data showed significant main effects of print, $F(1,60) = 6.22$, $MS_e = 0.045$, $p < 0.02$, and icons, $F(1,60) = 4.89$, $p < 0.04$. Comprehension of the warning message was significantly greater with conspicuous print ($M = 0.73$) than with plain print ($M = 0.60$) and greater when icons were present ($M = 0.72$) than when they were absent ($M = 0.61$). The interaction was not significant, $F(1,60) < 1.0$. The ANOVA on the strict data produced a similar pattern: significant main effects of print, $F(1,60) = 10.42$, $MS_e = 0.043$, $p < 0.01$, and icons, $F(1,60) = 4.66$, $p < 0.04$. Recall was greater with conspicuous print ($M = 0.57$) than with plain print ($M = 0.40$) and was greater when icons were present ($M = 0.54$) than when they were absent ($M = 0.43$). The interaction effect was not reliable, $F(1,60) = 2.06$, $p > 0.05$.

Icon identification. The icon identification means are shown in Table 1. Planned comparisons and the ANOVAs failed to find any significant effects in the liberally and strictly scored icon identification data. Test order was also examined, but it neither produced any reliable main effects nor interacted with the other variables.

Reliability. As in Experiment 1, 20 response sheets were randomly selected from both message content and identification tests and rescored by a second judge to assess reliability. For the message content test, the strict and liberal scores showed reliabilities of 0.96 and 0.86, respectively. For the icon identification test, the strict and liberal scores produced respective reliabilities of 0.97 and 0.92.

Discussion

As in Experiment 1, the results showed that conspicuous print, icons present warnings produced greater message content comprehension and memory than did the other three warning conditions. The ANOVAs on the liberal and strict scores showed that the print and icon factors additively influenced message content performance.

No effects were shown in the icon identification test. This was surprising given the results of Experiment 1 and because it was expected that exposure to the icons would (at least) produce greater icon identification accuracy than would no prior exposure to the icons in the manuals. As can be seen in Table 1, the liberal scores are rather high and the strict scores are rather low, suggesting that floor and ceiling effects might have limited the variability among scores, thus reducing the likelihood of finding differences between conditions.

GENERAL DISCUSSION

The two experiments demonstrate that instruction manual warnings that have both conspicuous print and illustrative pictorial icons enhance comprehension and memory of the warnings' message content. The presence of both dimensions of salience led to better performance than did the absence of one or both dimensions. Presumably the presence of conspicuous print and icons increased the warnings' noticeability (signaling quality) in a noisy background (the other text). Consequently, the added salience increased the likelihood that the warning messages would be read, comprehended, and remembered.

Another measure of comprehension and memory was assessed by the icon identification test, which examined whether icons can be used to cue retrieval of related hazard information. This would be applicable in situations in which a warning icon is seen only briefly and in which the icon alone must cue the user to the general and specific kind of hazard present. Experiment 1 showed that prior exposure to the icons facilitated general and specific hazard recall in the icon identification test. In addition, icons paired with conspicuously printed warning messages cued more information about specific haz-

ards than did icons paired with plainly printed warning messages. However, these icon identification effects were not confirmed in Experiment 2.

Because the two experiments were similar, some tentative comparisons across experiments may be useful; however, one should keep in mind that there are potential problems in interpreting apparent differences between experiments. As indicated in Table 1, the message content scores for the gas oven (Experiment 2) were generally greater than for the gas-powered electric generator (Experiment 1). The departure in performance might be attributable to differences in people's perceptions of the products (e.g., familiarity, perceived hazard, frequency of use). For example, familiarity with a product might facilitate encoding and/or retrieval of the associated safety information. Because the pattern of means was reasonably consistent between products, the message content results suggest that similar effects for print and icons might be found for warnings in instruction manuals for other products.

Examination of the liberal icon identification (comprehension) data across experiments showed approximately equivalent performance among conditions (except for the conspicuous print, icons present condition in Experiment 1). However, the strictly scored icon identification (memory) data showed better overall performance for the generator than for the oven. One possible reason for this difference is that participants spent less time processing the specific verbal messages associated with the oven icons. If this were the case, we would expect that message content performance would be lower for the oven than for the generator. However, as noted earlier, the data showed the opposite. Low performance on the oven's strictly scored icon identification test could be attributable to familiarity with other, similar oven products. This potential familiarity could lead to a failure in retrieval of specific information for this particular oven. At this point no definitive conclusion can be made regarding the different icon identification results between the two experiments.

Although previous research has noted positive effects of highlighting on reading warnings (Strawbridge, 1986) and compliance (Zlotnik, 1982), it has found little or no effect on comprehension and memory (Desaulniers, 1987; Otsubo, 1988; Rothstein, 1985; Strawbridge, 1986; Zlotnik, 1982). In addition, research has not shown significant effects of icons on warning comprehension and memory measures (Friedmann, 1988; Otsubo, 1988). However, the current study showed positive effects of conspicuous text and the presence of icons compared with less conspicuous text and the absence of icons on warning comprehension and memory. Recently DeJoy (1989) suggested that any of a number of methodological differences (e.g., different situations, stimulus materials, and instruction sets) among studies could be responsible for the discrepancies found in the warning literature. For example, unlike most previous work, the current research used a large number of warnings (eight) and presented them inside a manual (as opposed to on a product).

The enhancement shown by pairing the pictorial icons with verbal print has some theoretical support in the human memory literature. Paivio's (1975) dual-code theory asserts that a verbal code combined with an imaginal (pictorial) code leads to better memory than does either code alone, and a large body of basic cognitive research supports this notion (e.g., D'Agostino, O'Neill, and Paivio, 1977). Presumably both codes combine at the time of encoding to produce richer representations, which are subsequently easier to retrieve.

Two cautions should be mentioned regarding the interpretation of the present results. One concerns the print manipulation. Three

modes of salience were combined to produce the conspicuous print: fluorescent orange highlighting, different font style, and larger size. The current study does not separate the individual effects of these enhancement factors, nor does it indicate which factor was more influential. The positive effects shown by the conspicuous print manipulation points out that text enhancement factors are worthy of further study. For example, future research could examine the independent effect of each of these and other methods of print enhancement (e.g., spacing, color of the print, and background).

The other caution concerns the presentation duration. Considering that participants were given only 4 min to examine the 10-page manual, it is likely that the material was skimmed or browsed rather than systematically read and studied. Other presentation procedures could conceivably produce different results. For example, one might expect that with greater time allowed to study the instruction manual, many of the effects noted in the present report could disappear because performance in all conditions would be near perfect. However, the relatively short duration used in the current work is probably representative of the way consumers actually examine and use safety and owner's manuals. Many times if consumers read the literature at all, the material is merely scanned for pertinent or unusual information and only referred to when trouble arises. The current procedure corresponds to this nonsystematic use of instructional manuals. Thus when a short time is spent examining an instruction manual, warnings made more noticeable are more likely to be read.

Two major implications can be drawn from the present research. The first concerns the methodology. The techniques used in these studies may be useful in future research and testing of instruction manual warnings. We were able to assess what people gleaned from instruction manuals for potentially dangerous equipment without actually exposing them to dangers. That the ruse was believable was apparent from informal observations of participants' reactions before they were told they were in the "control" group. After reading the manuals, most participants rose from their chairs, apparently waiting for the experimenter to lead them out of the room to the equipment. In addition, during the debriefing period that followed, most participants said they thought they would be trying to operate the generator or the oven.

The second implication of the present research is its applicability and relevance for writers of instruction manuals. The results clearly show that the way in which warnings are presented is important: they should be salient relative to other background information (i.e., noticeable) and thus set the stage for further processing, including comprehension and subsequent recall of the warning message.

ACKNOWLEDGMENTS

The authors would like to thank Michael J. Kalsher and Kenneth R. Laughery, Sr., for their helpful comments on an earlier draft of this research. Portions of this article were presented at the 32nd Annual Meeting of the Human Factors Society (Young and Wogalter, 1988).

REFERENCES

Collins, B. L. (1983). *Use of hazard pictorials/symbols in the minerals industry* (Tech. Report NBSIR 83-2732). Washington, DC: U.S. Department of Commerce.

Collins, B. L., Lerner, N. D., and Pierman, B. C. (1982). *Symbols for industrial safety* (Tech. Report NBSIR 82-2485). Washington, DC: U.S. Department of Commerce.

Cunitz, R. J. (1981). Psychologically effective warnings. *Hazard Prevention, 17*, 5–7.

D'Agostino, P. R., O'Neill, B. J., and Paivio, A. (1977). Memory for pictures and words as a function of level processing: Depth or dual coding? *Memory and Cognition, 5*, 252–256.

DeJoy, D. M. (1989). Consumer product warnings: Review and analysis of effectiveness research. In *Proceedings of the Human Factors Society 33rd Annual Meeting* (pp. 936–940). Santa Monica, CA: Human Factors Society.

Desaulniers, D. R. (1987). Layout, organization, and the effectiveness of consumer product warnings. In *Proceedings of the Human Factors Society 31st Annual*

Meeting (pp. 56–60). Santa Monica, CA: Human Factors Society.

Dorris, A. L., and Purswell, J. L. (1978). Human factors in the design of effective product warnings. In *Proceedings of the Human Factors Society 22nd Annual Meeting* (pp. 343–346). Santa Monica, CA: Human Factors Society.

FMC Corp. (1985). *Product safety sign and label system.* Santa Clara, CA: Author.

Friedmann, K. (1988). The effect of adding symbols to written warning labels on user behavior and recall. *Human Factors, 30,* 507–515.

Godfrey, S. S., Allender, L., Laughery, K. R., and Smith, V. L. (1983). Warning messages: Will the consumer bother to look? In *Proceedings of the Human Factors Society 27th Annual Meeting* (pp. 950–954). Santa Monica, CA: Human Factors Society.

Laux, L., Mayer, D. L., and Thompson, N. B. (1989). Usefulness of symbols and pictorials to communicate hazard information. In *Interface 89: The Sixth Symposium on Human Factors and Industrial Design in Consumer Products* (pp. 79–83). Santa Monica, CA: Human Factors Society.

Lehto, M. R., and Miller, J. M. (1986). *Warnings: Fundamentals, design, and evaluation methodologies* (Vol. 1). Ann Arbor, MI: Fuller Technical Publications.

Mayer, D. L., and Laux, L. F. (1989). Recognizability and effectiveness of warning symbols and pictorials. In *Proceedings of the Human Factors Society 33rd Annual Meeting* (pp. 984–988). Santa Monica, CA: Human Factors Society.

McCarthy, R. L., Robinson, J. N., Finnegan, J. P., and Taylor, R. K. (1982). Warnings on consumer products: Objective criteria for their use. In *Proceedings of the Human Factors Society 26th Annual Meeting* (pp. 98–99). Santa Monica, CA: Human Factors Society.

Otsubo, S. M. (1988). A behavioral study of warning labels for consumer products: Perceived danger and use of pictographs. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 536–540). Santa Monica, CA: Human Factors Society.

Paivio, A. (1975). Perceptual comparisons through the mind's eye. *Memory and Cognition, 3,* 636–647.

Peters, G. A. (1984). A challenge to the safety profession. *Professional Safety, 29,* 46–50.

Robinett, F., and Hughes, A. (1984). Visual alerts to machinery hazards: A design case study. In R. Easterby and H. Zwaga (Eds.), *Information design: The design and evaluation of signs and printed material* (pp. 405–417). Chichester, England: Wiley.

Rothstein, P. R. (1985). Designing warnings to be read and remembered. In *Proceedings of the Human Factors Society 29th Annual Meeting* (pp. 684–688). Santa Monica, CA: Human Factors Society.

Strawbridge, J. A. (1986). The influence of position, highlighting, and imbedding on warning effectiveness. In *Proceedings of the Human Factors Society 30th Annual Meeting* (pp. 716–720). Santa Monica, CA: Human Factors Society.

Westinghouse. (1981). *Westinghouse product safety label handbook.* Trafford, PA: Westinghouse Printing Division.

Wogalter, M. S., Godfrey, S. S., Fontenelle, G. A., Desaulniers, D. R., Rothstein, P. R., and Laughery, K. R. (1987). Effectiveness of warnings. *Human Factors, 29,* 599–612.

Wright, P. (1981). The instructions clearly state . . . Can't people read? *Applied Ergonomics, 12,* 131–141.

Wright, P., Creighton, P., and Threlfall, S. M. (1982). Some factors determining when instructions will be read. *Ergonomics, 25,* 225–227.

Young, S. L., and Wogalter, M. S. (1988). Memory of instruction manual warnings: Effects of pictorial icons and conspicuous print. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 905–909). Santa Monica: Human Factors Society.

Zlotnik, M. A. (1982). The effects of warning message highlighting on novel assembly task performance. In *Proceedings of the Human Factors Society 26th Annual Meeting* (pp. 93–97). Santa Monica, CA: Human Factors Society.